

Are mutants really natural? A study on how “naturalness” helps mutant selection

Matthieu Jimenez¹, Thierry Titcheu Chekam¹, Maxime Cordy²,
Mike Papadakis¹, Marinos Kintis¹, Yves Le Traon¹, Mark Harman³

University of Luxembourg¹

University of Namur²

University College London and Facebook³

紹介担当者：本田 紘貴（大阪大学）

ミューテーションテスト

- テストスイートの十分さを測定するためのテスト
 - ① テスト対象のプログラムの一部を書き換える（ミュータント）
例：比較演算子>を<に置き換え
 - ② テストスイートをミュータントに対して実行
 - ③ テストスイートを元のプログラムに対して実行
 - ④ ②と③の結果が異なれば，テストスイートはその欠陥を発見するだけの能力（欠陥発見能力）があり
 - ⑤ ②と③結果が同じであれば，その欠陥を発見するためのテストを追加することでテストスイートの欠陥発見能力の向上可能

コードの自然さ

- コードの単語列が言語としてどれだけでもっともらしいか
 - コードにも自然言語のようにパターンや規範に従う傾向があり
 - この傾向からその言語の言語モデルを生成可能
 - 本論文では確率的言語モデルであるN-gram言語モデルを用いて、単語列のもっともらしさを数値で表現

概要

- 研究目的
 - ミュータントの自然さと欠陥発見との間に関連性はあるのか
 - 自然なミュータント（コーディング規範に従う自然な欠陥），または不自然な元のコード（一般的に欠陥が発生しやすい）から作られたミュータントは高い価値があるのではないかと筆者らは直観的に期待
- 調査方法
 - 元のコードの自然さ，ミュータントの自然さ，およびミュータントと元のコードの自然さの差異に対する欠陥発見率を調査し，関連を分析
- 結果(ネガティブ)
 - ミュータントの自然さと欠陥発見との間に関連性を発見することはできなかった

調査対象

- 欠陥と修正
 - Defects4Jの5つのプロジェクトにある230個の欠陥とその修正
- テスト
 - 自動テスト作成ツールで作成された1,375,341個とプロジェクト開発者が作成した58,131個のテストを利用
- ミュータント
 - ミュータント作成ツールMajor[21]を利用し, ミュータントを作成

調査方法(1/2)

言語モデルを用いて自然さを求め数値を算出

- (1) ミュータントの自然さ
- (2) 元のコードの自然さ
- (3) ミュータントと元のコードの自然さの差異（絶対値）

自然な場合と不自然な場合のどちらがより欠陥発見に関連するかを確認するために(1)(2)(3)をソートしセット分け

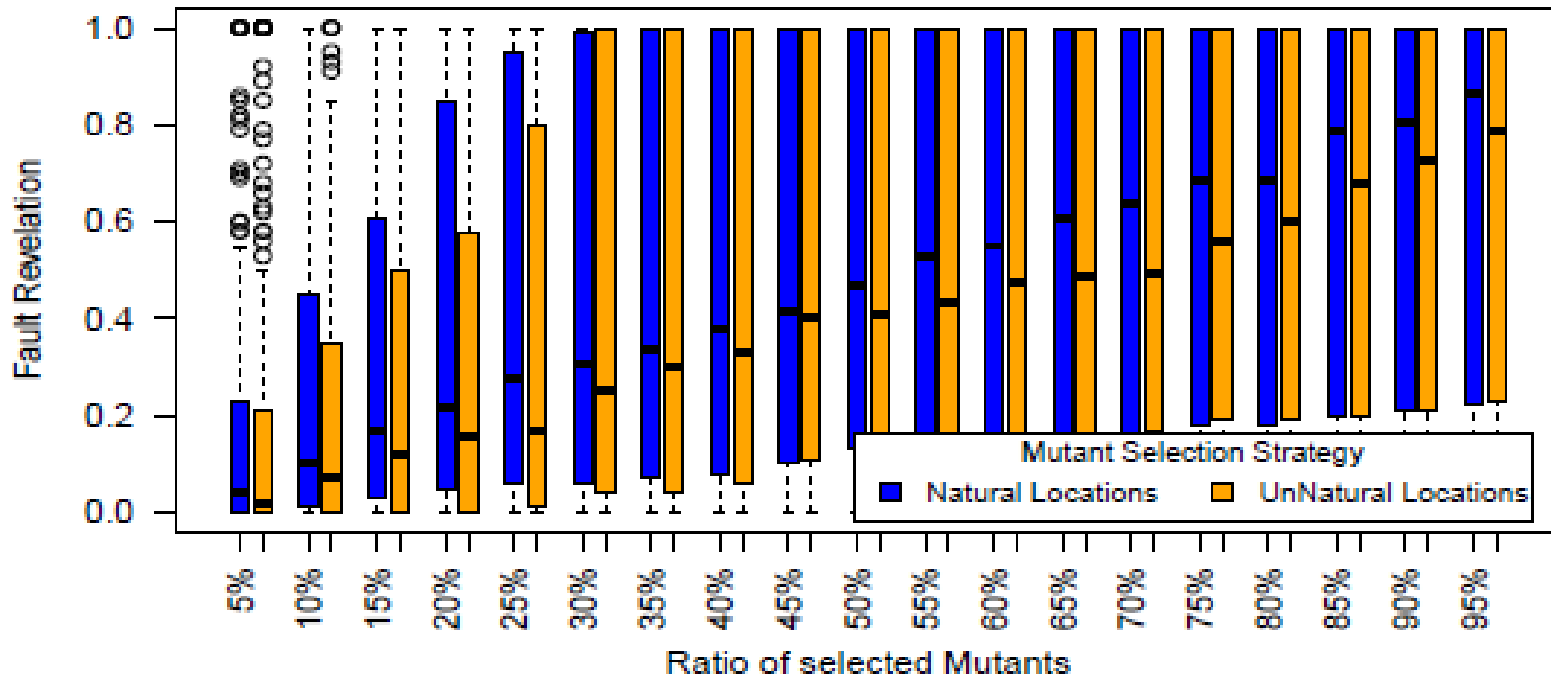
- 自然さが高いまたは低い順に上位（5,10,15,・・・,100%のセット）(1)と(2)
- 差異が大きいまたは小さい順に上位（5,10,15,・・・,100%のセット）(3)

調査方法(2/2)

- 各セットのミュータントにおいて、ミューテーションテストを適用
- 各セットのミュータントの欠陥をすべて検知するテストを選択
- 選択したテストを欠陥のあるコードに対して適用、実際のプロジェクトの欠陥を発見できるか調査
- 選択したテストのうち、実際のプロジェクトの欠陥を発見する率（欠陥発見率）を算出
- 各セットに対する欠陥発見率の関係を箱ひげ図を用いて図示

結果(1/3)

- 元のコードの自然さに対する欠陥発見率

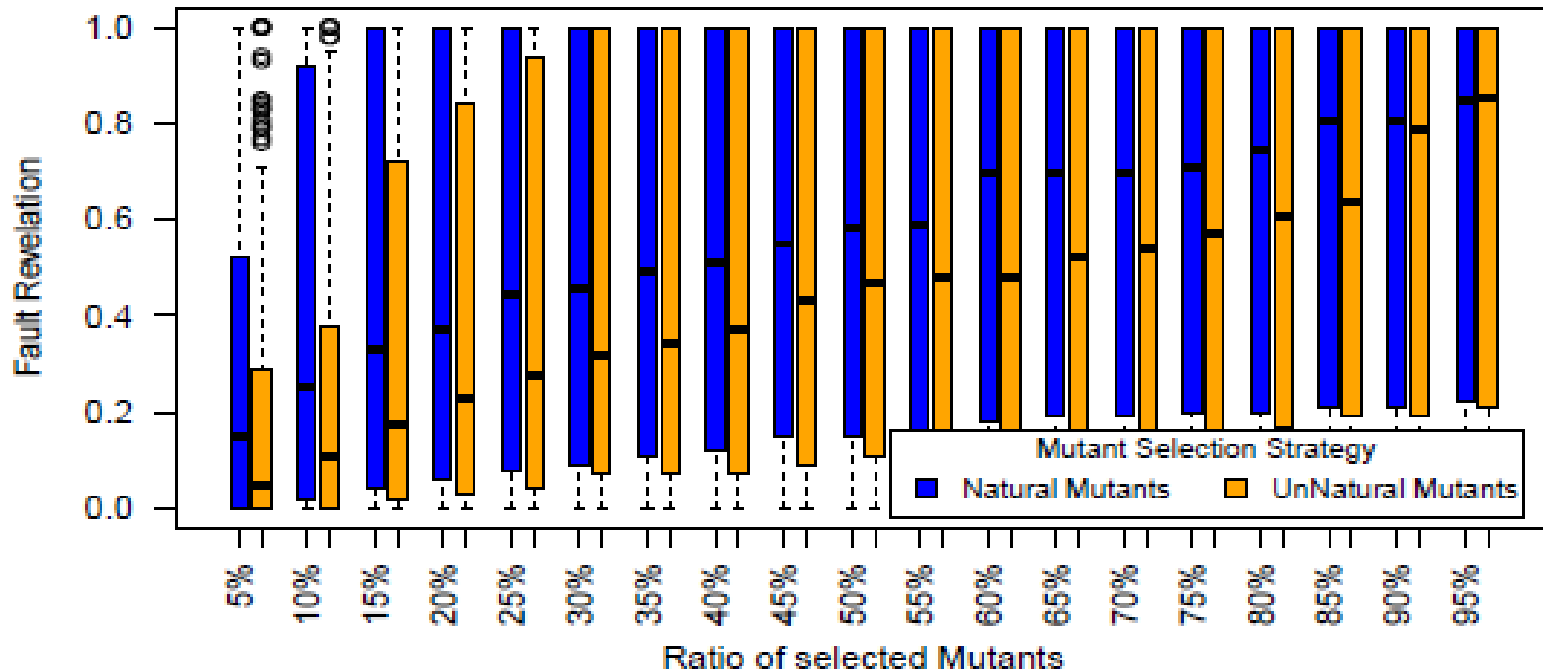


(a) Natural VS Unnatural code locations

Source: "Are mutants really natural? A study on how "naturalness" helps mutant selection" (ESEM2018) Fig.2 より一部抜粋

結果(2/3)

- ミュータントの自然さに対する欠陥発見率

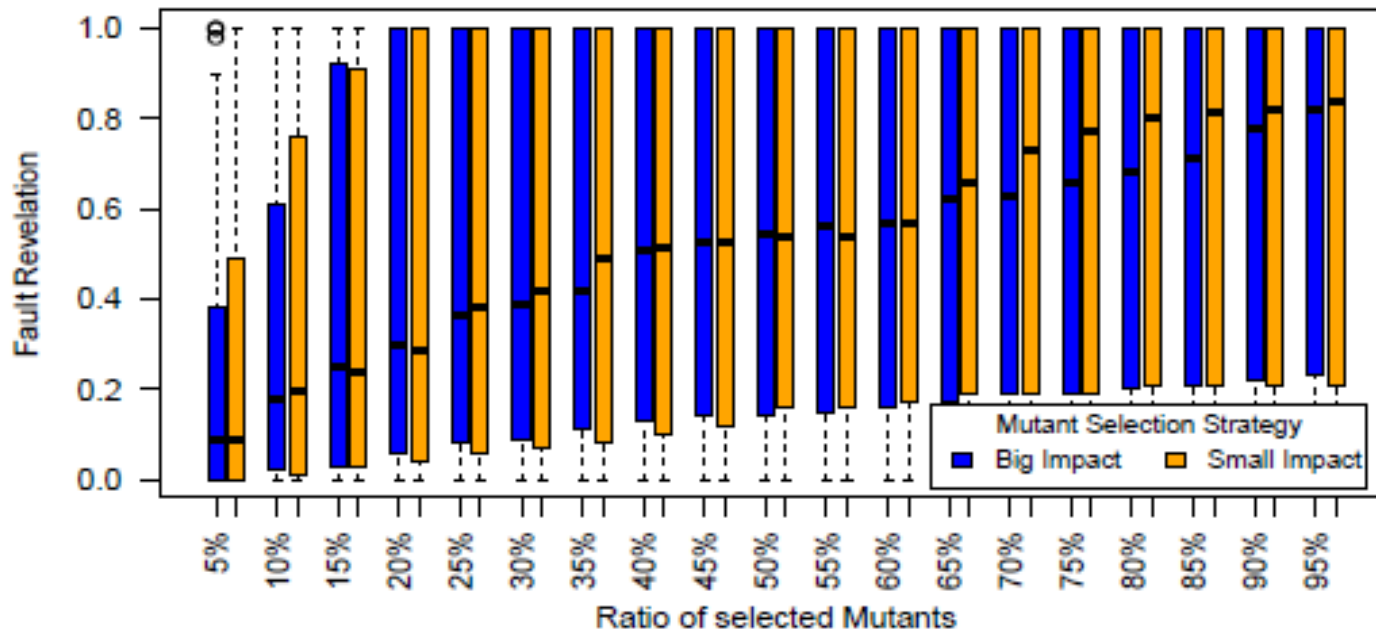


(b) Natural VS Unnatural mutants

Source: "Are mutants really natural? A study on how "naturalness" helps mutant selection" (ESEM2018) Fig.2 より一部抜粋

結果(3/3)

- ミュータントと元のコードの自然さの差異に対する欠陥発見率



(c) Big VS Small impact (entropy differences introduced by mutants)

Source: "Are mutants really natural? A study on how "naturalness" helps mutant selection" (ESEM2018) Fig.2 より一部抜粋

ミュータントと元のコードの自然さの差異と
欠陥発見に関連が見られない

まとめ

- 一般的に、不自然なコードは欠陥がある可能性が高い
 - 自然さに着目し、ミュータントの自然さと欠陥発見との関連を調査
- 言語モデルで自然さを算出し、欠陥発見との関連を分析
 - 元のコードの自然さ、ミュータントの自然さ、および元のコードとミュータントの自然さの差異に対する欠陥発見率を調べ、関連性を分析
- ミュータントの自然さと欠陥発見との間に**関連性を発見することはできなかった**